

JOURNAL Submitted 02-JUN-1999 Genoscope - Centre National de Séquençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>

FEATURES source

ORIGIN

Query Match 5.5%; Score 75.4; DB 29; Length 1101; Best Local Similarity 48.2%; Pred. No. 0.0084; Indels 0; Gaps 0; Matches 148; Conservative 10; Mismatches 149; Gaps 0;

QY 666 TTGAACTTATAATTACATAATTATAGTTCACTCTTATATACTCATAT 725
Db 183 TTTAAAWATWATAAAWAAATWATAAAATTAAWWAAAATAAAWAT 242

QY 726 AATATATAGGGTTAGATGCCAATTTTAAAAAGAATAAAATAAGATAAA 785
Db 243 AAAAAAWAAAWATAWWAAAATAAAATAAAATAAAATAAAATAAA 302

QY 786 ATCGAAATAATGAAATGTAAATAATTGAGGGACAATAATAATGAAAGTCATAT 845
Db 303 ATATATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 362

QY 846 TTAATTTCCATTAGATCTATTTCCTAGTTATGCTAGCTGGCAGAT 905
Db 363 TAATATTTTTTATTTTTATTTTTATTATAATAATAATAATAATAATAATA 422

QY 906 ACAGCAAATATGTATGAAACAGTGCATGTAGGAAATTAAATGTACTAGAGGATACCA 965
Db 423 AAAAAATAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 482

QY 966 GACAAAAA 972
Db 483 AAANAAA 489

RESULT 10 CNS0039G 1101 bp DNA linear GSS 03-JUN-1999

DEFINITION Drosophila melanogaster genome survey sequence TER3 end of BAC # BAC08K10 of RPCI-58 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AI063921

VERSION GI:4941778

KEYWORDS Drosophila melanogaster (fruit fly)

SOURCE Drosophila melanogaster

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidae; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)

AUTHORS Genoscope.

Db	422	AAASATGGWAAWATAAGHTTKDWAAAAGAANNNNNHCCCCCCCM@CCCCCTTY	363
Qy	854	TCCATTAGGATTCATTTCCCTAGTAAAT	885
Db	362	YYTTTTTTTTATMACTTACTTTAT	331

Search completed: February 24, 2004, 11:07:12
Job time : 4772.54 secs

RESULT 2
 US-10-257-166-114
 Sequence 114, Application US/10257166
 Publication No. US20040232301
 GENERAL INFORMATION:
 APPLICANT: OLEK, Alexander
 APPLICANT: PLEPENROCK, Christian
 APPLICANT: BERLIN, Kurt
 TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of Genes Implicated in Pharmacogenomics
 FILE REFERENCE: 5013.1011
 CURRENT APPLICATION NUMBER: US/10/257,166
 CURRENT FILING DATE: 2002-10-07
 PRIOR APPLICATION NUMBER: PCT/EP01/07470
 DE 10043826.1
 PRIOR FILING DATE: 2001-06-29
 2000-05-30
 2000-05-01
 SEQ ID NO 114
 LENGTH: 6106
 NUMBER OF SEQ ID NOS: 178
 SEQ ID NO 114
 LENGTH: 6106
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-257-166-114
 Query Match 4.7%; Score 64.4%; DB 16; Length 6106;
 Best Local Similarity 56.5%; Pred. No. 0; Indels 1; Gaps 1;
 Matches 139; Conservative 0; Mismatches 106;
 us-10-257-166-114

Query 1 TAAAATATCTACATTAaaaaaaatTCGTTAAATTTAGAAATTCTGATTTATTT 60
 Db 797 TATTTAA GTTTTATTAAGTTAATTTAGTTAAATTTAGAAATTCTGATTTATTT 856
 Qy 61 TTTCCTTACAGAAATCGGTTAAATATGTTAAATTTAGCAACTTAATAAATTTA 120
 Db 857 TTTCCTTATTTATTAATAATTTAGCAACTTAATAAATTTAATAAATTTA 916
 Qy 121 TTTCCTTACAGAAATCGGTTAAATATGTTAAATTTAGCAACTTAATAAATTTA 179
 Db 917 TTTCCTTATTTATTAATAATTTAGCAACTTAATAAATTTAATAAATTTA 976
 Qy 180 AGATAAAGTTGATGGCCGGGTGTTAGTTAGTTCTAGTTGAATGCAACT 239
 Db 977 ATATAATAATAGAAAGTATTTGTTATTTTATTTTATTTGTTGAACT 1036
 Qy 240 TGTTTT 245
 Db 1037 GAGTAT 1042

RESULT 3
 US-10-311-455-1196
 Sequence 1196, Application US/10311455
 Publication No. US20030143605A1
 GENERAL INFORMATION:

Qy 666 TTGAACTTATTATAATTACATATTATTAAGTTCACTTCTATAACTCATAT 725
 Db 200 TTAAANTTTAANANNTTTINAAAATTTTNNNNNNTTTAA.....AAAAAA 141
 Qy 726 AATATAGGGTTAGAATGCCAATTAAAGATAAAATAATAGAATAAA 785
 Db 140 AAAAANTGGNCNNNNNTTTTAA.....AAA.....AAA.....AAA.....AAA..... 81
 Qy 786 ATCGAAAAAATGAAATGTAAAAAATTGGGGCACAAATAATGAAAT 838
 Db 80 AAA.....AAA.....AAA.....ACCA.....AAA.....AAA.....AAA..... 28

RESULT 7
 US-09-814-353-12666/c
 Sequence 12666, Application US/09B14353
 Publication No. US2005016531A1
 GENERAL INFORMATION:
 APPLICANT: Lee, John
 APPLICANT: Thompson, Pamela
 APPLICANT: Lillie, James
 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
 FILE REFERENCE: MRI-0063
 CURRENT APPLICATION NUMBER: US/09/814,153
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: US 60/191,031
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: US 60/207,124
 PRIOR FILING DATE: 2000-05-25
 PRIOR APPLICATION NUMBER: US 60/211,940
 PRIOR FILING DATE: 2000-06-15
 PRIOR APPLICATION NUMBER: US 60/216,820
 PRIOR FILING DATE: 2000-07-07
 PRIOR APPLICATION NUMBER: US 60/220,661
 PRIOR FILING DATE: 2000-07-25
 PRIOR APPLICATION NUMBER: US 60/257,672
 PRIOR FILING DATE: 2000-12-21
 NUMBER OF SEQ ID NOS: 223/7
 SOFTWARE: FastQC for Windows Version 4.0
 SEQ ID NO: 12666
 LENGTH: 318
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 121..123, 124..126, 129..134, 157..158, 159..160, 161..
 LOCATION: 172..179, 180..183, 194..205, 210..213, 216..219, 220..
 LOCATION: 225..226, 227..235, 241..243, 248..250, 267..268, 272..277,
 LOCATION: 282..287, 293..302
 OTHER INFORMATION: n = A,T,C or G

US-09-814-353-12666

Query Match 4.6%; Score 63.2; DB 10; Length 318;
 Best Local Similarity 56.6%; Prod. No. 0.08;
 Mismatches 98; Conservate 0; Gaps 0

Qy 666 TTGAACTTATTATAATTACATATTATTAAGTTCACTTCTATAACTCATAT 725
 Db 200 TTAAANTTTAANANNTTTINAAAATTTTNNNNNNTTTAA.....AAAAAA 141
 Qy 726 AATATAGGGTTAGAATGCCAATTAAAGATAAAATAATAGAATAAA 785
 Db 140 AAAAANTGGNCNNNNNTTTTAA.....AAA.....AAA.....AAA..... 81
 Qy 786 ATCGAAAAAATGAAATGTAAAAAATTGGGGCACAAATAATGAAAT 838
 Db 80 AAA.....AAA.....AAA.....ACCA.....AAA.....AAA.....AAA..... 28

```

US-09-814-353-17949
Sequence 17949, Application US/09814353
Publication No. US20030165831A1
GENERAL INFORMATION:
APPLICANT: Lee, John, Pamela
APPLICANT: Thompson, Pamela
TITLE OF INVENTION: NOVEL GENES, COMPOUNDS AND METHODS FOR IDENTIFICATION, ASSESSMENT AND THERAPY OF OVARIAN CANCER
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/8141
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,941
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,824
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,677
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FastSEQ For Windows Version 4
SEQ ID NO: 17949
LENGTH: 403
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 225
OTHER INFORMATION: n = A,T,C or G
US-09-814-353-17949

Query Match          4.6%   Score: 67
Best Local Similarity 53.5%  Pred: Non-Mismatch
Matches 131; Conservative 0;
QY          603 TTGGATCTTGCTTGTGCGCTTTGTTGTTCTTC
Dy          115 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
QY          663 TACGTGAACCTATTAAATTAAATTACATAA
Dy          175 TTTTTTTTTTTTTTTTTTTTAAATAAATTT
QY          723 TATAATATAAGGGTTTAAATGCAATGAAAAAA
Dy          235 AAAAATAAATTTTATTGGATTAATAAAAAT
QY          783 AAAATCGAAAAAAATGAAATGTTAAAAAAAT
Dy          295 AAACCCATTAAAATAAAATAATTTC
QY          843 TATT 847
Dy          355 AATT 359

RESULT 9
US-09-960-352-4584
Sequence 4584, Application US/09960352
Patient No. US2002017139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nangbing C.
APPLICANT: Bratt, John C.
APPLICANT: Mathia Agan, Nagappan
TITLE OF INVENTION: MUSCLE ACID AND OXALATE LEVELS AND FAT DEPOSITION IN HUMAN MUSCLE TISSUE

```


Sequence 19050, Application US/09814353
 Publication No. US20030-65831A1
 GENERAL INFORMATION:
 APPLICANT: Lee, John
 APPLICANT: Thompson, Pamela
 APPLICANT: Tillye, James
 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER
 FILE REFERENCE: MR-L-006B
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: US 6/0791,031
 PRIORITY FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: US 6/207,124
 PRIORITY FILING DATE: 2000-05-25
 PRIOR APPLICATION NUMBER: US 6/211,940
 PRIORITY FILING DATE: 2000-06-15
 PRIOR APPLICATION NUMBER: US 6/0216,820
 PRIORITY FILING DATE: 2000-07-07
 PRIOR APPLICATION NUMBER: US 6/0220,661
 PRIORITY FILING DATE: 2000-07-25
 PRIOR FILING DATE: 2000-12-21
 NUMBER OF SEQ ID NOS: 22037
 SOFTWARE: Fast-SEQ for Windows Version 4.0
 SEQ ID NO: 19050
 LENGTH: 385
 TYPE: DNA
 ORGANISM: Homo sapiens
 SEQ ID NO: 09-814-353-12050
 Query Match 4.5%; Score 61; DB 10; Length 385;
 Best Local Similarity 59.5%; Pred. No. 0.21; Indels 0; Gaps 0;
 Matches 103; Conservative
 22Y 666 TTGAAACTTATTATTAATTTACATAATTATTAATGTTTCACTCTTAT 725
 Ddb 268 TTAAATTTTAAATAAAATTTAAATAAAATTGGTTTTAAAAAAA 209
 22Y 726 AATATAAGGTAGATGCCATTTCATAAAAGATAAAATAATAGATAAA 785
 Ddb 208 AAAAATGGGGCATTTTTAAAAAATTTAAATAATGAAAGT 838
 22Y 786 ATCGAAATAAAATTAATAAAATTGAGGGGACAAATAATAATGAAAGT 838
 Ddb 148 AAAAATAAAATAAAATCAAATAAAATAAAATAAAATAAAAGT 96
 RESULT 13
 Sequence 10-311-455-8133-8133/C
 Publication No. US200311455
 GENERAL INFORMATION:
 APPLICANT: OLBK, Alexander
 APPLICANT: PIEPENBROCK, Christian
 APPLICANT: BERLIN, Kurt
 TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ FILE REFERENCE: 5013-1014
 CURRENT FILING DATE: 2002-12-16
 PRIOR APPLICATION NUMBER: PCT/EP01/07537
 PRIORITY FILING DATE: 2001-07-02
 PRIOR APPLICATION NUMBER: DE 100322529-7
 PRIORITY FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: DE 10043826-1
 NUMBER OF SEQ ID NOS: 2424
 LENGTH: 7631
 TYPE: DNA
 ORGANISM: Artificial Sequence
 SEQ ID NO: 833
 Query Match 4.4%; Score 60.2; DB 12; Length 6487;
 Best Local Similarity 50.7%; Pred. No. 0.87; Indels 2; Gaps 1;
 Matches 170; Conservative
 US-10-221-613-227
 11 TATACATTAAAAATTGATTAAATTGAAATTGATTATTTTACGTTTAC 70
 Ddb 4081 TATACATTTTTATTTGATATAAGTTAAATTTTAAAGGGCA 4140
 11 AGAAATCCGTTAAATGTTAAATGTTAACTTATTTATTTTAAAGGGCA 130
 Ddb 4141 AGTTAAAGTTTAAAGATAATTAGTAAATTATTTATTTGAG 4200
 131 ATTACGATTTGATATAATATATGTTAAATAACTTCA-AGAGTAAAG 188
 Ddb 4201 TATTAAATTGATGTTGATATAAGTTAAATTTAAAGTTAGGT 248
 189 TTGATGATGCGGTGTTAGATTTAGTCTGAGTCGAGTGGTTTGA 4261
 Ddb 4261 ATAATAGCTTTATTTGATATAAGTTAAATTTAAAGGTTATA 4320

Qy 249 CATTAGCCCTATCCTTCTAACCAAAATAATGAAATGGAAACCTTAGGAAA 308
 Db 4321 GTTTTTTAAATGGCTTTTGTTCATAATAGAAATGCTGGATAGTACGG 4380

Qy 309 AAACAAATAAATGAAACATATCCCGTGA 313
 Db 4381 AATGATTAATGAAACAAATAAATAGGAA 4415

RESULT 15

US-10-311-455-1594

Sequence 1594, Application US10311455

GENERAL INFORMATION:

APPLICANT: OLEK, Alexander

APPLICANT: PIRPENBROCK, Christian

APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of Invention: Cytosine methylation

CURRENT APPLICATION NUMBER: US10/311,455

CURRENT FILING DATE: 2002-12-16

PRIOR APPLICATION NUMBER: PCT/EP01/07537

PRIOR FILING DATE: 2001-07-02

PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: DE 10043826.1

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 2424

SEQ ID NO: 1594

LENGTH: 6254

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1594

OTHER INFORMATION:

Query Match 9 4.4%; Score 60; DB 14; Length 6254;

Best Local Similarity 53.9%; Pred. No. 0.93;

Matches 166; Conservative 0; Mismatches 140; Indels 2; Gaps 2;

Db 2641 TTTCATATATAGTTGTTTAAATTTAAAGAATTTTTAAAGTA 2700

Qy 69 CCAGAAATCGTTAAATGTTAAATACCAACTAAATTAAATTTAAAG 128

Db 2701 TAAGGATAAAATTAATTTAAATAGATAAAATTTTTATAGATT 2760

Qy 129 CAATTAGCATGTTGATAAATATATATTGTTAAATACTTTCAAATAG 188

Db 2761 AAACAAATAATTG-TATTTAAATATAATTTAAATATTTATT 2619

Qy 189 TTGATGATGGGTGGTGGATATTAA-TGTTAGTTGAAATGCAAGTCTTAG 247

Db 2820 TTGGATGATGATTATTGTTATAATATTTTTTTTGTTAGAAATGTTTA 2879

Qy 248 ACATTTGCCATTCTTCTAACCAAAATAATGAAACCTTAGGAA 307

Db 2880 TTAGGTAATATTTATTTGAACTGTTAAATGAGGATGAGTGANGT 2939

Qy 308 AAAAGAA 315

Db 2940 ATAAGAA 2947

BEST AVAILABLE COPY

US-09-083-587-3

DESCRIPTION: /desc = "mouse PAL cDNA"

Query Match Score 55; DB 4; Length 2246;
Best Local Similarity 57.1%; Pred. No. 0.042;
Matches 100; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

662 GTACTTGAACATTATAATCATATTATAACTCTTAAATTAATATC 721
2049 GTATTGAAACAGCTTCAAGTAGCTTCAATTTGGTCATGGCT 2108

722 ATATPATATATGGCTTAAAGATSCCAATTTTAAGATAAAATAATGAA 781
2109 GTGTATTATTAAGTGTACTGACCAGAAAAAAATGAA 2168

782 TAAATCTGAAATGAAATGAAATGAAATGAAATGAA 836
2169 AAAAAGAAAAAAAAGAAAAAAAGAAAAAAAGAAAAAA 2223

RESULT 5
SEQUENCE 1 Application US/09790988
Patent No. 663235

GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEMI
APPLICANT: HATTORI, MASAHIRU
APPLICANT: SAKAKI, YOSHIOUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL Symbiont OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SEQUENCE ID NO 1
SEQUENCE 2 Application US/09790988
Patent No. 663235

Query Match Score 54.8; DB 4; Length 640681;
Best Local Similarity 42.9%; Pred. No. 0.061;
Matches 381; Conservative 0; Mismatches 502; Indels 5; Gaps 2;

1 TAAATAATCTATACATTAAATTATGTTAAATTAGATA 60
353537 TTACTAATPATCATTAGCTTAAATCTTAAATAGATA 353478

61 TTTTTTTACCGAATCCCTTAATATGTTAAATTAGATA--ATTATAATT 117
353477 AATTTTTTTATAAATCTTAAATCTTAAACAAATAGCTATAGAA 353418

118 TATTTTAAGCAATTAAAGATGTGTTGATAAAATATATTTAAATCTCA 177
353417 ATCAATTACGGTAATGCAATTATATTTAAATAAATTTAA 353358

178 AAGGTATAAAGTTGATGATGCCGTGGTAGATTATGTTCTAGTTGATGCA 237
353357 GAATATTTCGTTGATGATCATTAAAGTATCTTAAATATAGTA 353298

238 GTTGGTTAGACATTAGCTTATCTTCTAACCAAATATGAAATGGAAAC 297
353297 AGGATTTAAATAAACATTATTTAAATTTAAATTATCTT 353238

298 CTTAGGAAARAAAAGAAATCAAAATTGAAACATCATCGGTGAGAGCCAC 357
353237 TCTCGGTATTATCTAAATGATAACAAAARCCTTAATTTAAATTATAG 353178

358 ACCACGTGACCCACAAATATAAGTTGTTCTACGATA 417

RESULT 6
US-07-867-106-2/C
Sequence 2 Application US/07867106
Patent No. 5389526

GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
APPLICANT: Improved Plasmid Vectors for Cellular
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & NO. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 1992-06-25
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: AU FJ 7187
APPLICATION NUMBER: PCPAU90/00530
FILING DATE: 0-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFEECTION DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 5852 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ANTI SENSE: NO
 FEATURE: CDS
 NAME KEY: CDS
 LOCATION: 2378..5038
 NAME KEY: CDS
 LOCATION: 2378..5038
 US-07-867-106-2

Query Match 4.0%; Score 54.21; DB 1; Length 5852;
 Best Local Similarity 58.6%; Fred. No. 0.038; Indels 1; Gaps 1;
 Matches 112; Conservative 0; Mismatches 78; Db 5713

```
Qy 683 TTACATATTATAAGTTTCACTTCUTATA-TAATACCATATAATATAGGGTTAG 741
Db 5713 TRACTTATTCTTATTATTAAATTAAATTAAATTAAATTAAATTAAATAAAA 5654
```

Query Match 4.0%; Score 54.21; DB 1; Length 615;
 Best Local Similarity 60.1%; Fred. No. 0.034; Indels 2; Gaps 1;
 Matches 107; Conservative 0; Mismatches 69; Db 26

```
Qy 5 ATATATCATACATTAAAAATTGATTTAGAATTCTATGATTTATTTTT 64
Db 26 ATAAGGATAATATACCTTTTATATAAGTAAATTAAACTTAAACTTAAACTT 85
```

Query Match 3.9%; Score 53.6%; DB 3; Length 615;
 Best Local Similarity 60.1%; Fred. No. 0.034; Indels 2; Gaps 1;
 Matches 107; Conservative 0; Mismatches 69; Db 65

```
Qy 65 TTACCGAACAAACCGTAA-TTGTAAATTATCCACTAAATTATAATTTTTT 122
Db 86 ATTATCATTATTAAATTAAATTAAATTATGATTATACTTAAATTAAATTATA 145
```

Query Match 3.9%; Score 53.6%; DB 3; Length 615;
 Best Local Similarity 60.1%; Fred. No. 0.034; Indels 2; Gaps 1;
 Matches 107; Conservative 0; Mismatches 69; Db 123

```
Qy 123 TTAGGGATTAACTGGCAATTGATAAAATATGTAATTTCAAACTTTCAAA 180
Db 146 TAATTTCTTAATTCACTTAAATTATATAATTAAATAAAATAATTATAA 203
```

RESULT 8
 US-08-998-416-1137
 Sequence 1137, Application US/08998416

PATENT INFORMATION:
 Patent No. 6239264

GENERAL INFORMATION:
 APPLICANT: Philippson, Peter
 APPLICANT: Pohlmann, Rainer
 APPLICANT: Steiner, Sabine
 APPLICANT: Mohr, Christine
 APPLICANT: Wendland, Jürgen
 APPLICANT: Knechtle, Philipp
 APPLICANT: Reibischung, Corinne
 TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHYA GOSSYPII
 NUMBER OF SEQUENCES: 1152
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 6239264artis Corporation
 STREET: 3054 Cornwallis Road
 CITY: Research Triangle Park
 STATE: No. 6239264th Carolina
 COUNTRY: USA
 ZIP: 27709
 ZIP: 27709
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ParentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/998,416
 FILING DATE: 24-DEC-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: CH 0016/97
 FILING DATE: 31-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEX: 919-541-8689

Dbs 782 TATATTTTAAATAAATTAACATCTC 756

RESULT 11
US-08-213-419B-3
Sequence 3, Application US/08213419B
Patent No. 6333406
GENERAL INFORMATION:
APPLICANT: Inselburg, J. et al.
TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: JII-002CNCP
CURRENT APPLICATION NUMBER: US/08/213,419B
CURRENT FILING DATE: 1994-03-14
PRIOR APPLICATION NUMBER: US 07/870,506
PRIOR FILING DATE: 1992-04-17
NUMBER OF SEQ ID NOS: 20
SEQ ID NO 3 Patent in Ver. 2.0
TYPE: DNA
ORGANISM: Plasmodium falciparum
FEATURE:
NAME/KEY: CDS
LOCATION: (2407) .. (2419)
NAME/KEY: CDS
LOCATION: (2598) .. (3404)
NAME/KEY: CDS
LOCATION: (3580) .. (3720)
NAME/KEY: CDS
LOCATION: (3850) .. (5825)
US-08-213-419B-3

Query Match 3.8%; Score 52.6; DB 4; Length 6124;
Best Local Similarity 52.0%; Pred. No. 0.079;
Matches 118; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 710 TATATAATCATATAATATAATAGGGTTAGAAATGCCATTAAAGGATAAAA 769
Db 2089 TAATAAGGGCCATATCTCAATTCTCTTAAGATGTACAAAAA 2148

Qy 770 AAATAATAGATAAAATGAAAAATTGAAATGTAAAATTGAGGGGACAAATAAA 829
Db 2149 AAATTAATAAATAAAATAAAATAATTTAAATATTTATTTATTTAA 2208

Qy 830 TATGAAGTCTATTATTAATTTCCATTAGAACATCTTCCATTAGTTATGAGC 889
Db 2209 TTATATATATATATAATATAATTTAAATTTAAATTTAAATACATGTCAAA 2268

Qy 890 TAGCCAGTTGGAGATACAGAAAATGTCATGAAACACTGTCATGA 936
Db 2269 AAAAATAAAATAAAATAATATAATTTAAATATAATACATGTCATATA 2315

RESULT 12
US-08-998-416-515
Sequence 535, Application US/08998416
GENERAL INFORMATION:
APPLICANT: Philippson, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knecht, Philipp
APPLICANT: Reibischung, Corinna
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHYA GOSSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264arts Corporation
STREET: 3054 Cornwallis Road

Query Match 3.8%; Score 52.2; DB 3; Length 827;
Best Local Similarity 49.5%; Pred. No. 0.068;
Matches 135; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 24 ATTGGATTTAAATTAGAAATTCTGATTTTATCCAGAAATCCGTTAA 83
Db 194 AATTGATTTATCTTAAACATAAAACATTAAATATAAAATAAAT 253

Qy 84 TATTGTAATAAAATTACAACTPAATTTAAATTAAAGGAAATTAGCATGTT 143
Db 254 TACTTAAAGAAATTATTTAAATAGTATTAAATTTAAATATAATACATT 313

Qy 144 GATAAAATATATAATTGTTAAATCTTCAAAGGATAAAGTTGATGGGTGG 203
Db 314 TTATTAATATAATGATTAAATAGTTTAAATATGTCATAATAATTTATATA 373

Qy 204 TGGTAGATTATTAGTTCTAGTTCTGAATGTTAGACATTAGCCATTTC 263
Db 374 AAATTATTTAACTTCATGATAATAATTTAAATATAATTTAAATGACCT 433

Qy 264 TTTTCTAACAAATAAACTGAAATGAAA 296
Db 434 ATTTTATTAGTCATGAAATTTCTTAAATAA 466

RESULT 13
US-09-641-638-651
Sequence 651, Application US/09641638
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marla
APPLICANT: Bougueret, Lydia
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knecht, Philipp
APPLICANT: Reibischung, Corinna
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264arts Corporation
STREET: 3054 Cornwallis Road

PRIOR APPLICATION NUMBER: US 09/502,330
 PRIORITY FILING DATE: 2000-02-11 : deletion of C
 PRIORITY FILING DATE: US 60/133,200
 PRIORITY FILING NUMBER: US 1999-05-07
 PRIORITY FILING DATE: 1999-03-23 : variable motif ATTA or TTTT
 PRIORITY FILING NUMBER: US 60/119,917
 PRIORITY FILING DATE: 1999-02-12 : polymorphic base C or T
 NUMBER OF SEQ ID NOS: 1304
 SOFTWARE: Patent-PM

SEQ ID NO: 651
 LENGTH: 20674
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE: misc_feature
 LOCATION: 1123..3133
 OTHER INFORMATION: 5' regulatory region
 NAME/KEY: exon
 LOCATION: 3124..3237
 OTHER INFORMATION: exon 1
 NAME/KEY: exon
 LOCATION: 3871..4072
 OTHER INFORMATION: exon 2
 NAME/KEY: exon
 LOCATION: 5552..5633
 OTHER INFORMATION: exon 3
 NAME/KEY: exon
 LOCATION: 5758..5880
 OTHER INFORMATION: exon 4
 NAME/KEY: exon
 LOCATION: 5996..6039
 OTHER INFORMATION: exon 5
 NAME/KEY: exon
 LOCATION: 6340..6509
 OTHER INFORMATION: exon 6
 NAME/KEY: exon
 LOCATION: 7379..7522
 OTHER INFORMATION: exon 7
 NAME/KEY: exon
 LOCATION: 8645..8854
 OTHER INFORMATION: exon 8
 NAME/KEY: exon
 LOCATION: 12254..12340
 OTHER INFORMATION: exon 9
 NAME/KEY: exon
 LOCATION: 12854..13023
 OTHER INFORMATION: exon 10
 NAME/KEY: exon
 LOCATION: 13308..13429
 OTHER INFORMATION: exon 11
 NAME/KEY: exon
 LOCATION: 16567..16667
 OTHER INFORMATION: exon 12
 NAME/KEY: exon
 LOCATION: 16775..16945
 OTHER INFORMATION: exon 13
 NAME/KEY: exon
 LOCATION: 17063..17554
 OTHER INFORMATION: exon 14
 NAME/KEY: misc_feature
 LOCATION: 17555..20674
 OTHER INFORMATION: 3' regulatory region
 NAME/KEY: allele
 LOCATION: 1128
 OTHER INFORMATION: 10-508-191 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 1182
 OTHER INFORMATION: 10-508-245 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 1559
 OTHER INFORMATION: 10-509-284 : polymorphic base C or T
 NAME/KEY: allele

LOCATION: 1570
 OTHER INFORMATION: 10-509-295 : deletion of C
 NAME/KEY: allele
 LOCATION: 1827
 OTHER INFORMATION: 10-510-173 : variable motif ATTA or TTTT
 NAME/KEY: allele
 LOCATION: 2048
 OTHER INFORMATION: 10-511-62 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 2323
 OTHER INFORMATION: 10-511-337 : insertion of T
 NAME/KEY: allele
 LOCATION: 2341
 OTHER INFORMATION: 10-512-36 : polymorphic base G or C
 NAME/KEY: allele
 LOCATION: 2623
 OTHER INFORMATION: 10-512-318 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 283
 OTHER INFORMATION: 10-513-250 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 2844
 OTHER INFORMATION: 10-513-262 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 2934
 OTHER INFORMATION: 10-513-352 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 2947
 OTHER INFORMATION: 10-513-365 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 3802
 OTHER INFORMATION: 12-206-81 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 4062
 OTHER INFORMATION: 10-513-365 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 4088
 OTHER INFORMATION: 12-206-366 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 4109
 OTHER INFORMATION: 10-343-231 : deletion of C
 NAME/KEY: allele
 LOCATION: 4170
 OTHER INFORMATION: 10-343-339 : polymorphic base G or T
 NAME/KEY: allele
 LOCATION: 5903
 OTHER INFORMATION: 10-343-278 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 6019
 OTHER INFORMATION: 10-346-141 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 6141
 OTHER INFORMATION: 10-346-263 : polymorphic base G or C
 NAME/KEY: allele
 LOCATION: 6183
 OTHER INFORMATION: 10-346-305 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 6338
 OTHER INFORMATION: 10-347-74 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 6375
 OTHER INFORMATION: 10-347-111 : polymorphic base G or C
 NAME/KEY: allele
 LOCATION: 6429
 OTHER INFORMATION: 10-347-165 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 6467
 OTHER INFORMATION: 10-347-203 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 6484
 OTHER INFORMATION: 10-347-220 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 6534

ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,055A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchell, Leonard C.
REGISTRATION NUMBER: 23,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-1200
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
S-08-446-055A-1

Query Match	Score	DB 2;	Length
Best Local Similarity	3.8%;	Score 51.6;	
Best Local Similarity	53.5%;	Pred. No. 0.13;	
Matches 108;	Conservative	Indels 0;	Gaps 0
Matches 108;	Conservative	Indels 0;	Gaps 0
662	GTACTGAACTTATTAATTACATAATTATTAAGTTTCACTCTTATATAATCTC	T 721	
526	GTCCTCATATAATTATTAATTATTAAGTTTCACTCTTATATAATCTC	T 585	
722	ATATATATATAGGTTTGTAGATGCCATTITTAALARAGATAAAAAATATAGAA	T 781	
586	TTAGTTTATATAATATGTTAATCTCTATTTAALARAGATAAAAAATATAGAA	T 645	
782	TAATATCGAAATATGAAATATGAAATATGAGGGGACAAATAAATAGAACTCTA	T 841	
646	GRAAAALAAAATATGAAATATGAAATATGAAATATGAACTGTATATGTAATTT	T 705	
842	TTATTTAAATTTCCATTAGA	T 863	
706	AAACATAAATATAATCTATA	T 727	

Search completed: February 24, 2004, 11:09:40
Job time : 138.614 secs

TEST AVAILABLE COPY

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 06:53:45 ; Search time 705.304 Seconds

Perfect score: 8233.739 Million cell updates/sec

Title: US-03-743-823-21

Sequence: ttaaaataatctataattaa.....ctcacacactagaaata 1367

Scoring table: IDENTITY_NUC

GapOp 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124039041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : N_Geneseq_297an04:**

1: Geneseqn1980s:**
 2: Geneseqn1980s:**
 3: Geneseqn2000s:**
 4: Geneseqn2001as:**
 5: Geneseqn2001bs:**
 6: Geneseqn2002s:**
 7: Geneseqn2003as:**
 8: Geneseqn2003bs:**
 9: Geneseqn2003cs:**
 10: Geneseqn2004s:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description	
1	1367	100.0	1367	3 AAZ58125	Aaz58125 Soybean 7	
2	1282.8	93.8	1428	9 ACP80573	Acf80579 Basic 7S	
3	74.6	5.5	8056	7 ABZ10246	Abz10246 Haemotopo	
C	4	72.4	5.3	8056	7 ABZ10246 Haemotopo	
5	71.4	5.2	8056	7 ABZ10100	Abz1100 Haemotopo	
6	64.4	4.7	6106	4 AAS46430	Aas46430 Tumour su	
7	64.4	4.7	6106	6 ABK40012	Abk40032 Human che	
8	64.4	4.7	6106	6 ABZ13473	Abz13473 Human imm	
9	64.2	4.7	7509	6 ABZ13223	Abz13223 Human imm	
10	64	4.7	309	5 ABV44994	Abv44994 Human pro	
C	11	63.4	4.6	7195	4 AAS45325	Aas45325 Chemical
C	12	63.4	4.6	7195	6 ABK28166	Abk28166 DNA trans
C	13	63.2	4.6	416	7 ABX39419	Abx39419 Bovine ES
C	14	63.2	4.6	8056	7 ABZ10100	Abz1100 Haemotopo
C	15	62.2	4.6	375	7 ABX49849	Abx49849 Bovine ES
C	16	62	4.5	11218	7 ACR64845	Aca6845 Human HNR
C	17	61.2	4.5	626	5 ABY60941	Abv60941 Human pro
C	18	61.2	4.5	551	6 ABZ34001	Abz34001 Human imm
C	19	60.4	4.4	7631	6 ABL32860	Abl32860 Human imm
C	20	60.2	4.4	6487	6 AAS41362	Aas41362 Signal tr
C	21	60.2	4.4	6487	6 ABZ10319	Abz10319 Chemical
C	22	60	4.4	6487	6 AAS61266	Aas61266 Human gen
C	23	60	4.4	6254	6 ABZ13621	Abz13621 Human imm

ALIGNMENTS

RESULT 1
 AAZ58125 standard; DNA, 1367 BP.

ID AAZ58125
 XX AC AAZ58125/
 XX DT 25-APR-2000 (first entry)
 XX DE Soybean 7S basic globulin gene promoter.
 XX KW Globulin; promoter; soybean; lactoferrin; human; transgenic Plant;
 KW seed storage protein; ds;
 XX OS Glycine max.

XX PN WO200004146-A1.

XX PD 27-JAN-2000.

XX PP 19-JUL-1999; 99WO-IT000226.
 XX PR 17-JUL-1998; 9BIT-RN000478.
 XX PR 17-JUL-1998; 9BIT-RN000478.

(PLAN-) PLANTECHNO SRL.
 PA (PLAN-) PLANTECHNO SRL.
 XX PI Fogher C;
 XX DR WPI; 2000-161129/14.
 XX PT Synthetic polynucleotide encoding human lactoferrin, used for production of functional foods, vegetal milks and human lactoferrin.
 XX PS Claim 7; Page 85-86; 93pp; English.
 XX CC The present sequence is that of the promoter region of the soybean var.
 CC Michland Seed-specific 7S basic Globulin gene. The Promoter region is incorporated into an expression cassette of the invention also including a synthetic gene (see AAZ58122) encoding human lactoferrin and designed for expression of human lactoferrin in transgenic plants. Use of particular leader sequences and promoters allows expression of the CC lactoferrin in specific tissues, especially seeds. The transgenic plants CC can be used in processes for the production of functional vegetal milk, CC fruit juices, fruit and/or vegetable homogenized foods (claimed). CC Expression of lactoferrin in plants should overcome problems associated CC with microbial expression and simplify purification. Expression in seeds CC also leads to an increase in iron content of seed products. The CC

		Sequence	1367 BP;	475 A;	245 C;	202 G;	445 T;	0 U;	0 Other;
CC	CC	Query Match	100.0%	Score 1367;	DB 3;	Length 1367;			
CC	CC	Best Local Similarity	100.0%	Pred. No. 6.4e-212					
CC	CC	Matches 1367;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps			
CC	CC								
CC	CC	1	TAAATAATCTACATTAAAGATTGATTTTAAATTTAGAAATTCTAGATTATTATT						
CC	CC	1	TAAATAATCTACATTAAAGATTGATTTTAAATTTAGAAATTCTAGATTATTATT						
CC	CC	61	TTTTTACCGAAAATCGGTTAATATGTAAATAATTAACTAAATTATAAATTATA						
CC	CC	61	TTTTTACCGAAAATCGGTTAATATGTAAATAATTAACTAAATTATAAATTATA						
CC	CC	121	TTTAAGCCAAATTAAGCATGTTGATTAATAATTCTTAAATTTCTAGTTGAACT						
CC	CC	121	TTTAAGCCAAATTAAGCATGTTCAATAATAATTATGTTAAATTTCTAGTTGAACT						
CC	CC	181	GTATAAAGTGTATGATGGCCGGCTGGTGTTAGATTATTTCTAGTTGAACT						
CC	CC	181	GTATAAAGTGTATGATGGCCGGCTGGTGTTAGATTATTTCTAGTTGAACT						
CC	CC	241	GGTTAACATTAGCCTTATCTTCTTCTTAACCAAATAAACTTAATGAAACCTT						
CC	CC	241	GGTTAACATTAGCCTTATCTTCTTCTTAACCAAATAAACTTAATGAAACCTT						
CC	CC	301	TAGAAAAAAGAAATCAAAATGAAACATCCGGGACTTCGAAGAGCCACCC						
CC	CC	301	TAGAAAAAAGAAATCAAAATGAAACATCCGGGACTTCGAAGAGCCACCC						
CC	CC	361	CACGTGACCACAAATTAAATAAGACTTGCCTACGTAATTGCTAACTTTTTA						
CC	CC	361	CACGTGACCACAAATTAAATAAGACTTGCCTACGTAATTGCTAACTTTTTA						
CC	CC	421	TCCAATCTTTCACCTCTAAACCTTGCTGAGATTGCCCCCTAACATTAAAGTGT						
CC	CC	421	TCAATCTTTCACCTCTAAACCTTGCTGAGATTGCCCCCTAACATTAAAGTGT						
CC	CC	481	AATGCCACGGCCTTAAACACTTGCTACCGTGCCTCACATTCACTTTGCCACCC						
CC	CC	481	AATGCCACGGCCTTAAACACTTGCTACCGTGCCTCACATTCACTTTGCCACCC						
CC	CC	541	TGAAGCGGTATGGTTAGTTAGCTAGCTGTTGAAACAGTTGAGGGCATCTCAGGGGT						
CC	CC	541	TGAAGCGGTATGGTTAGCTAGCTGTTGAAACAGTTGAGGGCATCTCAGGGGT						
CC	CC	601	ACTGGCTCTGCCTTCTGGCTTTGGTCATTTCACCGCATTTGGTGAACAGT						
CC	CC	601	ACTGGCTCTGCCTTCTGGCTTTGGTCATTTCACCGCATTTGGTGAACAGT						
CC	CC	661	CGTACTGAAACCTTAATTACATAATTTTATAAGTTTCACTCTTCTATAAATCT						
CC	CC	661	CGTACTGAAACCTTAATTACATAATTTTATAAGTTTCACTCTTCTATAAATCT						
CC	CC	721	CATAATAATAAGGGTTAGATGCCATTTCCTTAGTTAAATAGGTTAGGGAT						
CC	CC	721	CATAATAATAAGGGTTAGATGCCATTTCCTTAGTTAAATAGGTTAGGGAT						
CC	CC	781	ATAAAATGGAAAAATGAAATGTAAMAAATTGAGGGACAAATAATATGAAAGTCT						
CC	CC	781	ATAAAATGGAAAAATGAAATGTAAMAAATTGAGGGACAAATAATATGAAAGTCT						
CC	CC	841	ATTTTAATTTCCATTAGAATCTTATTCCTTAGTTAAATAGGTTAGGGAT						
CC	CC	841	ATTTTAATTTCCATTAGAATCTTATTCCTTAGTTAAATAGGTTAGGGAT						
CC	CC	901	GAGATACAGAAATGTCATGAAACTGCTGTTGGAAATTATGTTAGGGAT						
CC	CC	901	GAGATACAGAAATGTCATGAAACTGCTGTTGGAAATTATGTTAGGGAT						

active and in a high amount suitable for medical use

Query Match 4.7%; Score 64.4; DB 6; Length 6106;
 Best Local Similarity 56.5%; Pkd. No. 0.13; Mismatches 0;保守性 1; Indels 1; Gaps 1;
 Matches 139; Conservatve 0; Mismatches 106; Indels 1; Gaps 1;

Qy 1 TAAATATCTACATAAAATTGATTTGAAATTGATTTGAAATTGATTTT 60
 Db 797 TAATTAAGTTTATTTAGAGTTTAAAAATCGGTAAATTGTTTT 856
 Qy 61 TTTCCTTACCGAAATCCGTTAATATGTTAAATACCAATTATAAATTAA 120
 Db 857 TTTCCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 916
 Qy 121 TTTCAGGAAATTAAGCATGGTT GATAAAATATATATGTTAAATTGTTAA 179
 Db 917 ATTAATTGATGTTAAATTGTTAGATTAAGATTAAGATGTTGTA 976
 Qy 180 AGTATAAATCTGATGATGCCGTTGCTAGTTCTAGGTCTAGGTCTAGGT 239
 Db 977 AATATTAATTAGAAGTAGTTGTTATTGTTATTGTTATTGTTGTTGAGT 1036
 Qy 240 TGGTT 245
 Db 1037 GAGTAT 1042

RESULT 9
 ABL33223
 ID ABL33223 standard, DNA, 7560 BP.
 XX DE Human immune system associated gene SEQ ID NO: 1196.
 XX DT 26-MAR-2002 (First entry)
 XX DE Human immune system associated gene SEQ ID NO: 1196.
 KW Human; immune disease; cytosine methylation; antiallergic;
 KW antiarteriosclerotic; antianæmic; cytostatic; nonotropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antiinflammatory; anticancer; antidiabetic; antipsoriatic;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 KW ds.
 XX Homo sapiens.
 XX OS WO200200928-A2.
 XX PN 03 -JAN-2002.
 XX PP 02-JUL-2001; 2001WO-EP007537.
 XX PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIG-) EPIGENOMICS AG.
 XX P1 Olek A, Piepenbrock C, Berlin K,
 XX DR, 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 XX PS Claim 1, SEQ ID NO 1196, 32BP + Sequence listing: German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX Sequence 7560 BP; 2154 A; 47 C; 1554 G; 3805 T; 0 U; 0 Other;

Query Match 4.7%; Score 64.2; DB 6; Length 7560;
 Best Local Similarity 50.5%; Pres. No. 0.14; Mismatches 0; Conservatve 156; Indels 0; Gaps 0;
 Matches 156; Conservatve 0; Mismatches 153; Indels 0; Gaps 0;

Qy 1 TAAATATCTATACATAAAATTGATTTGAAATTGATTTGAAATTGATTTT 60
 Db 2072 TAAATTAAGTAACTGAAATTGATTATTTTATTTTATTTTGTGTTTATT 2131
 Qy 61 TTTCCTTACCGAAATTCGGTTAATTTGTTAAATATACCAACTAAATTATAAATTAA 120
 Db 2132 ATTITATTGTTAAAGTGTAAATTGTTAGATTAAGATTAAGATGTTGTA 2191
 Qy 121 TTTCAGGAAATTAGCAGTTGATAAAATATATTTGTTAAATTTTCAAAA 180
 Db 2192 GTTGTAGATTAGAAGTAGTTAAAGATTTAAATTTTAAATGTTAAATAT 2251
 Qy 181 GTATAAAGTGTGATGGCGTGGTGTGATTTGTTAGGTCTAGGTGCGACTT 240
 Db 2252 TTTCATTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 2311
 Qy 241 CGTTAGACATTAGCCTTATTCCTTCTAACAAATAAACTGAATGGAAACCTT 300
 Db 2312 GATTGTGAAATTAAATTAAATTAAATTAAATTAAATTAAATTAAATTAA 2371

RESULT 10
 ABV44994
 ID ABV44994 standard, cDNA; 309 BP.
 XX AC ABV44994;
 XX DT 16-SEP-2002 (First entry)
 XX DE Human prostate expression marker cDNA 44985.
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX DD 23-AUG-2001.
 XX EP 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0183862P.
 PR 25-MAY-2000; 2000US-020754P.
 PR 09-JUN-2000; 2000US-021314P.
 PR 18-JUL-2000; 2000US-0215007P.
 PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PA Schlegel R, Endgege WO, McNahan JB;
 XX PI WPL; 2001-662795/76.

XX DR
 XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 CC Claim 1; Page 8916; 11750pp; English.
 CC
 CC

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABY001-ABY6213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 SQ Sequence 309 BP; 188 A; 13 C; 11 G; 97 T; 0 U; 0 Other;
 Query Match 4.7%; Score 64; DB 5; Length 309;
 Best Local Similarity 56.0%; Pred. No. 0.14;
 Matches 121; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 Qy 621 CTTTGTTGAAATTTCAGTGATTCTTGGTGAACACCGTACTGAAACTTATAA 680
 Db 19 CTTTTTTTTTTTATATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 78

Qy 681 AATTACATATTATTTATAAGTTTCACTTCTTATATACTCATATAATAGGGTTTA 740
 Db 79 TTA 138
 Qy 741 GAATGCCAAATTTCATTTAAAARGATAAAATAATAGATAAAATCGAAAAATGAA 800
 Db 139 AAAAATTTGAAAAAAATTTAAAAAAATTTAAAAAAATTTAAAAAAATTTAAA 198
 Qy 801 TGTTAAATTTTGAGGGCAGCAATAATATGAA 836
 Db 199 AAAAATTTAAAAAAATTTAAAAAAATTTAAAAAAATTTAAAAAAATTTAAA 234

RESULT 11
 AAS45325/C
 ID AAS45325 standard; DNA; 7195 BP.
 AC AAS45325;
 XX DT 18-DEC-2001 (first entry)
 XX DE Chemically pretreated complementary DNA associated with cell cycle #15.
 XX Cell cycle: human; CpG dinucleotide; cytosine methylation; HIV; aging;
 XX human immunodeficiency virus; neurodegenerative disorder; solid tumor;
 XX arthritis; arteriosclerosis; graft-versus-host disease; glomerular disease; cancer;
 XX immunosuppressive; antitumour; cytostatic; antiarthritic; PCR primer.
 XX OS Homo sapiens.
 XX PN WO200168911-A2.
 XX PD 20-SEP-2001.
 XX PP 15-MAR-2001; 2001WC-EP002945.

XX PR 15-MAR-2000; 2000DE-01013547.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032329.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIC-) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K;
 XX DR 2001-602751/6.

XX Designing primers and probes for analyzing diseases associated with PT cytosine methylation state e.g. arthritis, cancer, aging, arteriosclerosis comprising fragments of chemically modified genes associated with cell cycle.

XX PS Claim 1; SEQ ID NO 30; 28pp; English.

XX Sequences AAS45296 AAS45520 represent chemically pretreated genomic DNA molecules associated with the cell cycle and specific PCR primers of the CC invention. The sequences are useful for detecting the methylation state CC of all CpG dinucleotides in a sequence and therefore for analysing CC associated diseases. By analysing cytosine methylations in the pretreated CC DNA, Genetic and/or epigenetic parameters for the diagnosis and therapy CC of existing diseases or the predisposition to specific diseases can be CC ascertained. The parameters may be compared to another set of genetic CC and/or epigenetic parameters, the differences serving as basis for CC diagnosis and/or prognosis events which are disadvantageous to patients. CC The sequences of the invention are useful for the diagnosis and therapy CC of HIV infection, neurodegenerative disorders, graft-versus-host disease, CC aging, glomerular disease, Lewy body disease, arthritis, CC arteriosclerosis, solid tumours and cancers

XX SQ Sequence 7195 BP; 1873 A; 224 C; 1538 G; 3560 T; 0 U; 0 Other;

Query Match 4.6%; Score 63.4%; DB 4; Length 7195;

Best Local Similarity 50.2%; Pred. No. 0.18;

Matches 157; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

Qy 675 ATTATAATTACATATTATTTAGTTTCACTTCTTATAATCTCATATAATATAG 734
 Db 6759 ATTATCCAAAACATCAATTAAATTTTAAACCATATTTTAACTTTAAATT 6700

Qy 735 GGTTTGAATGCCCCATTTTAAAGAGATAAAATAAGATAAAATCGAAAAAA 794
 Db 6699 TATTTAAATAATAAAACTAAATAAACATCTAAATTAAATAAAATAAAATAAA 6640

Qy 795 ATGAAATGTTAAATTTGGGGGACAATAATAATGAAAGTCATTTAAATTTT 854
 Db 6639 AAAACATTAAATTTAAATTTCAAAAACAAAATTTAAATTTAAACATCTAAAC 6580

Qy 855 CCATGAAATTCTTATTTCCTTATTATGAGCTAGCAGTGGGAGTAGAGAAA 914
 Db 6579 CCAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAACACAAACA 6520

Qy 915 TGTCTGAAACAGTGTGATGTTGGAAATAAATGTTAGAGGGATGAAAGCAAAAT 974
 Db 6519 ACTACAGATCATAATAATCATATTTTACATATACTACATACATACATAC 6460

RESULT 12
 ABK28166/C
 ID ABK28166 standard; DNA; 7195 BP.
 XX AC ABK28166;
 XX DT 23-APR-2002 (first entry)
 XX DE DNA transcription associated complementary genomic DNA #420.
 XX PR 15-MAR-2000; 2000DE-01013547.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032329.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIC-) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K;
 XX DR 2001-602751/6.

XX AC ABK28166;
 XX DT 23-APR-2002 (first entry)
 XX ID ABK28166 standard; DNA; 7195 BP.
 XX AC ABK28166;
 XX AC ABK28166;
 XX AC ABK28166;

KW polyglutamine disorder; solid tumour.
 XX
 Unidentified.
 OS
 WO200192265-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 06-APR-2001; 2001W0-EP003973.
 XX
 PR 06-APR-2000; 2000DB-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-BEP-2000; 2000DB-01043826.
 XX
 PA (EPICG-) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-09046/12.
 XX
 PT New nucleic acids or oligomers, useful for diagnosing or treating diseases associated with DNA transcription, e.g. immunological disorders, Werner syndrome, poriasis, myocardial infarction, solid tumors or cancer.
 PT
 PS Claim 1; SEQ ID NO 40; 32pp; English.
 XX
 CC The invention relates to a nucleic acid, which comprises a segment of the chemically pretreated DNA of genes associated with DNA transcription from one of 345 sequences, and an oligomer, in particular an oligonucleotide or peptide nucleic acid (PNA)-oligomer that hybridizes to or is identical to the chemically pretreated DNA of genes associated with DNA transcription. The set of oligomer probes are useful for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) in a chemically pretreated genomic DNA. The nucleic acids are useful for diagnosing or treating diseases associated with DNA transcription (particularly with the methylation status), e.g. adenosine deaminase deficiency, viral infection, retroviral infection, Sezary syndrome, haematological disorders, immunological disorders, developmental disorders, neurological disorders, psoriasis, Werner's syndrome, tuberculosis, neurodegenerative disorders, Waardenburg syndrome, Niemann-Pick disease, mealyodysplastic syndrome, myocardial infarction, hypertension, angiogenesis, erythropoiesis, congenital heart disease, HDR syndrome, arthritides, polyglutamine disorders, solid tumors or cancer. Sequences ABK28127-ABK28472 represent DNA transcription associated Genomic DNA molecules of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from the European Patent Office.
 CC Sequence 7195 BP; 1873 A; 224 C; 1538 G; 3560 T; 0 U; 0 Other;
 SQ Query Match Score 4.6%; Score 63.4%; DB 6; Length 7195;
 Best Local Similarity 50.2%; Pred. No. 0.18;
 Matches 157; Conservative 0; Mismatches 156; Indels 0;
 QY 675 ATTATAATTCACATTAATTTCATAACTTCTCATTATACTCATATATAG 734
 DB 6759 ATTATCCAAACATCATTTTAACTATTTAACATTTAATTATTATT 6700
 QY 735 CGTTAGAAGGCCATTTCATTAATGAAATGAAATCGAAAAA 794
 DB 6699 TATTAAAATATAAACATCTAAATAAAATAAACATCTAAATATAA 6640
 QY 795 ATGAAATGTAAAATTTGGGGACAAATAATGAAAGCTTAAATT 854
 DB 6639 AAAACATTTATAATTTCACAAACAAAAAATTTAAATACATTAAC 6580
 QY 855 CGATTAGAAATCTTATTTCTTAGTTAATTGAGCTTACGCTAGTGGAGATACGAAA 914
 DB 6579 CCAAAATTAAAATTTAAATCCATCAAATATAAAACACAA 6520

KW 915 TGTCTATGAAACAGTGTGAGTGTAGGAAATTAAATGAGAGGATACGACAGCAAAAT 974
 DB 6519 ACTAAAGTAAATAAAATACATATATTTACATATAATATTTACTATATAAA 6460
 PN 975 CCAAGCCAACTTA 987
 DB 6459 ACCCAACAAATA 6447

RESULT 13
 ABX3919
 ID ABX3919 standard; cDNA; 416 BP.
 XX
 AC ABX39419;
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DS Bovine EST associated with lactation/muscle/fat deposition #4584.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; IMFD/
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW Gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 PR 24-SEP-2001; 2001US-00960352.
 XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 PA (BYATT J C.; MATHLAGAN N.
 PA (TARONI) TDO N.
 PA (WARRIN W C.).
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 DR WPI; 2003-110599/10.
 XX
 PT New nucleic acid associated with lactation, and muscle and fat deposition, useful for genome mapping, gene identification and analysis, cattle breeding, or for genetically improving cattle.
 XX
 PS Claim 2; SEQ ID NO 4584; 245pp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated IMFD), derived from cattle, and the IMFD nucleic acid can specifically hybridise to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ABX34836-ABX49947, or complements of them. Also included are (1) a transformed cell having a nucleic acid comprising an IMFD nucleic acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The IMFD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine IMFD EST (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in

PR 11-JAN-2000; 2000US-00480902.

XX (BYAT/) BYATT J C.
 PA (MATH/) MATHILAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX

PI Byatt JC, Mathilagan N, Tao N, Warren WC;

XX DR WPI: 2003-110599/10.

XX PT New nucleic acid associated with lactation, and muscle and fat deposition, useful for genome mapping, gene identification and analysis, cattle breeding, or for genetically improving cattle.
 XX

PS SEQ ID NO 15014; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMFD), derived from cattle, and the LMFD nucleic acid can specifically hybridise to second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ABX34836-ABX4947, or complements of them. Also included are (1) a transformed cell having a nucleic acid comprising an LMFD nucleic acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMFD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine LMFD SSR (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site:

seqdata.uspto.gov/sequence.html?DocID=20020137139

XX Sequence 375 BP; 143 A; 14 C; 24 G; 194 T; 0 U; 0 Other;

Query Match Score 4.6%; Score 62.2; DB 7; Length 375;

Best Local Similarity 53.0%; Pred. No. 0.28; Matches 133; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 619 GTCTTGTGCAATTTCACCGTGAATTGGTGAACAGCTACTTAACTTATA 678

Db 276 GTATTTTTAATTTTTTTTTTCCATCCTTTTAAAAAAA 217

Qy 679 TAATTCACATTATAAGTTTCACTCTTATAATCTCATATAATAGGGTT 739

Db 216 TATATTTTTTTTATTTTAAATAATAATAATAATAATAATAATAATA 157

Qy 739 TAGATGCCATTAAATAAAAGATAAAATAATAGAATAATACTGAAATAATGA 798

Db 156 AATAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATA 97

Qy 799 ATGTAATAAATTTGAGGGACATAATAATAATAATAATAATAATAATAATA 858

Db 96 AAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 37

Qy 859 TAGAATCTAT B69

Db 36 AAAAATTTAT 26

TITLE A synthetic polynucleotide coding for human lactoferrin, vectors, cells and transgenic plants containing it
JOURNAL WO 0004146-A-27-JAN-2000;
FOGHER CORSAIR (IT); PLANTECHNO SRL (IT)

FEATURES Location/Qualifiers

source 1..1367 /organism="Glycine max"
/mol type="unassigned DNA"
/db_xref="Ct exon_3847"
1..1367 /note="sequence of the promoter of the gene coding for 7S basic globulin"

ORIGIN

Query Match 100.0%; Score 1367; DB 6; Length 1367;
Best Local Similarity 100.0%; Pred. No. 2.6e-202;
Matches 1367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAAATAATCATTAATAAATTGATTTAATTTAGAATTCAAGATTTATT 60
Db 1 TAAATAATCATTAATAAATTGATTTAATTTAGAATTCAAGATTTATT 60
Qy 61 TTTCCTTACCGAATAATCCGTAAATATGTTAAATAATTAAATTTA 120
Db 61 TTTCCTTACCGAATAATCCGTAAATATGTTAAATAATTAAATTTA 120
Qy 121 TTAAAGGCAATAAGCATGGTTGATAAAATAATATTGTTAAATAA 180
Db 121 TTAAAGGCAATTAGCATGGTTGATAAAATAATATTGTTAAATAA 180
Qy 181 GTAAAGATTGATGATGGTGTGGTGTGATATTAGTTAGTTAATT 240
Db 181 GTAAAGATTGATGATGGTGTGGTGTGATATTAGTTAGTTAATT 240
Qy 241 GGTTAGACATTAGCCATTACCTTAAACCAATAATGAAACCTT 300
Db 241 GGTTAGACATTAGCCATTACCTTAAACCAATAATGAAACCTT 300
Qy 301 TAGAAAAAAAATAATGAAACATAATGAAACTATCGGGTGAAGTCGACCC 360
Db 301 TAGAAAAAAAATAATGAAACATAATGAAACTATCGGGTGAAGTCGACCC 360
Qy 361 CACGGACCCAAACATAATGAAACATAATGAAACTATCGGGTGAAGTCGACCC 360
Db 361 CACGGACCCAAACATAATGAAACATAATGAAACTATCGGGTGAAGTCGACCC 360
Qy 421 TTCAACATTTTCCACTCTAAATCTGGAGATTGCACTTTAACCTTAAGTT 480
Db 421 TTCAACATTTTCCACTCTAAATCTGGAGATTGCACTTTAACCTTAAGTT 480
Qy 481 ATATCCAACGGTCTAAAAAACCTTGTAATGCTGCCTCAATTGCACTTGACCC 540
Db 481 ATATCCAACGGTCTAAAAAACCTTGTAATGCTGCCTCAATTGCACTTGACCC 540
Qy 541 TGAAGGCCCTATGTTTGGCTTCAATTGCACTTGACCTGAGGGCT 600
Db 541 TGAAGGCCCTATGTTTGGCTTCAATTGCACTTGACCTGAGGGCT 600
Qy 601 ACTGGCTCTGGCTTCAATTGCACTTGACCTGAGGGCTTGTCGACACG 660
Db 601 ACTGGCTCTGGCTTCAATTGCACTTGACCTGAGGGCTTGTCGACACG 660
Qy 661 CGTACTGAACTTATAATTACATAATTAAAGTTCACTCTTATAACT 720
Db 661 CGTACTGAACTTATAATTACATAATTAAAGTTCACTCTTATAACT 720
Qy 721 CATAATAATAGCTTGTGCAATTTTTAAAGATAATAATAAGATA 780
Db 721 CATAATAATAGCTTGTGCAATTTTTAAAGATAATAATAAGATA 780
Qy 781 ATAAATCCAAAATAATGAAATGTAATAATTGAGGGGACAATAATAAGCT 840
Db 781 ATAAATCCAAAATAATGAAATGTAATAATTGAGGGGACAATAATAAGCT 840
Qy 781 ATAAATCCAAAATAATGAAATGTAATAATTGAGGGGACAATAATAAGCT 840

RESULT 2

AX840232 AX840232 1428 bp DNA linear PAT 16-DEC-2003

LOCUS AX840232 Sequence 6 from Patent WO0307839.

DEFINITION Sequence AX840232

ACCESSION AX840232.1 GI:39978632

VERSION AX840232.1

KEYWORDS

SOURCE Glycine max (soybean)

ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytina; Magnoliophyta; eudicots; eudicots; Fabales; Papilionoideae; Phasoleae;
rosids; eurosids I; Fabaceae; Glycine.

REFERENCE Fogher, C. and Reggi, S.
Expression of lysosomal enzymes in plant seeds
Patent: WO 0307839-A 6.12-SEP-2003;
Location/Qualifiers

AUTHORS Fogher, C. and Reggi, S.
TITLE Expression of lysosomal enzymes in plant seeds
JOURNAL Patent: WO 0307839-A 6.12-SEP-2003;
FEATURES Location/Qualifiers

ORIGIN

Query Match 93.8%; Score 1282.8; DB 6; Length 1428;
Best Local Similarity 95.6%; Pred. No. 2.8e-189;
Matches 1365; Conservative 0; Mismatches 2; Indels 61; Gaps 2;

Qy 1 TAAATAATCATTAATAATTCAATTAAATTGAAATTCTGTTTATT 60
Db 1 TAAATAATCATTAATAATTCAATTAAATTGAAATTCTGTTTATT 60
Qy 61 TTTCCTTACCGAATAATCCGTAAATATGTTAAATAATTAAATTTA 120

Db	61	TTTTTTACCGAGAATCCGTTAATATTGTTAAATATTCCAACATAATTATAATTTTA	120	Qy	1140	TTTATCCTTCCTTGCCTCTGGAAATTCAAGCTCAAAAGCTCCAAATCTT	1199
Qy	121	TTTAAGGCAATTAGCATTTGTATAAAATAATAATATGTTATAAAATCTTCAA	180	Db	1201	TTTATCCTTCCTTGCCTCTGGAAATTCAAGCTCAAAAGCTCCAAATCTT	1260
Db	121	TTTAAGGCAATTAGCATTTGTATAAAATAATAATATGTTATAAAATCTTCAA	180	Qy	1200	GTCACCAACACGCCAAGAAAGAAATCTCATCCGTTACACTAGCCACTTAAG	1259
Qy	181	GTATAAGTGTATGATGGCCGGTGGTGTAGTTCTAGGGTGGAAATGCAAGTT	240	Db	1261	GTCACCAACACGCCAAGAAAGAAATCTCATCCGTTACACTAGCCACTTAAG	1320
Db	181	GTATAAGTGTATGATGGCCGGTGGTGTAGTTCTAGGGTGGAAATGCAAGTT	240	Qy	1260	CTAGGCCAACGGGTGATCTTCCTATAATATGTTAGCTCTTAACACAAACACTAACCA	1319
Qy	241	GGTTTAGACATTAGCCTTAATCTTTTCAACAAAATAATGTTATGGAAACCTT	300	Db	1321	CTAGGCCAACGGGTGATCTTCCTATAATATGTTAGCTCTTAACACAAACACTAACCA	1380
Db	241	GGTTAGACATTAGCCTTAATCTTTTCAACAAAATAATGTTATGGAAACCTT	300	Qy	1320	TATTCCATATTCACACTTGCTCTAACTACACACTAGAGATA	1367
Qy	301	TAGGAAAAGAAGAATCAAATTGAAATACTATCGGGGGAGTGAGACACTC	360	Db	1381	TATTCCATATTCACACTTGCTCTAACTACACACTAGAGATA	1428
Db	301	TAGGAAAAGAAGAATCAAATTGAAACATCATCCGGGGAGTGAGACACTC	360				
Qy	361	CACTGACCAACAAATTAAATAATAGAGTTGCGCTCACAGTAAATCCSATACTTTTA	420				
Db	361	CACTGACCAACAAATTAAATAATAGAGTTGCGCTCACAGTAAATCCSATACTTTA	420				
Qy	421	TTCAAATCACTTTCCACTCTCTAAATCTTGGAGATTGCAACCCTTAACATAATTAGTGT	480				
Db	421	TTCAAATCACTTTCCACTCTCTAAATCTTGGAGATTGCAACCCTTAACATAATTAGTGT	480				
Qy	481	ATATCCOLACGGTCCPRAAAAACCTGTGACCTGCTCCATTCAACTTGGCACCC	540				
Db	481	ATATCCAAACGGTCCPRAAAAACCTGTGACCTGCTCCATTCAACTTGGCACCC	540				
Qy	541	TGAAGCCGT-TATGTTAGTGTAGCTTGCACAGTGAAGCCATCACTCAGGAGC	599				
Db	541	TGAAGCCGTCTATGTTAGTGTAGCTTGCACAGTGAAGCCATCACTCAGGAGC	600				
Qy	600	TAGTGTGCTGCTTGGCCTTGGCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG	659				
Db	601	TAGTGTGCTGCTTGGCCTTGGCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG	660				
Qy	660	GGCTACTGAACTATTATAATTACATAATTATAAACTTCACTCTTATAAAATAC	719				
Db	661	GGCTACTGAACTATTATAATTACATAATTATAAACTTCACTCTTATAAAATAC	720				
Qy	720	719				
Db	721	TTCATTCATGGATTATAATTGTGATAATAAAACAGTTGTTAAAAATAATTAT	780				
Qy	720	TCTATATATATAAGGTTAGAATGCCATTTTAAAGATAAAAAAAATAATG	779				
Db	781	TCTATATATATAAGGTTAGAATGCCATTTTAAAGATAAAAAAAATAATG	840				
Qy	780	AATAAAATGAAAGAAATGAAAGTAAATGAAATAAAATAAAATAAAATGTC	839				
Db	941	AATAAAATGAAAGAAATGAAAGTAAATGAAATAAAATAAAATAAAATGTC	900				
Qy	840	TATTATTAAGAAATTCTCATAGAAATTCTTCCCTAGTTATGACCTAGCCAGTG	899				
Db	901	TATTATTAAGAAATTCTCATAGAAATTCTTCCCTAGTTATGACCTAGCCAGTG	960				
Qy	900	GGAGATACGAAATTCTCATAGAAATTCTCATAGAAATTCTCATAGAAATTCT	959				
Db	961	GGAGATACGAAATTCTCATAGAAATTCTCATAGAAATTCTCATAGAAATTCT	1000				
Qy	960	TAGCAAGACAAAAATCCAAAGCCAAAGTCTAGCTGCTCACGCAACTGATCCACAGTCTT	1019				
Db	1021	TAGCAAGACAAAAATCCAAAGCCAAAGTCTAGCTGCTCACGCAACTGATCCACAGTCTT	1080				
Qy	1020	TACAGGTTCAAAGGATGAATCTGCAACTAAACGCTTCTCACTG	1079				
Db	1081	TACAGGTTCAAAGGATGAATCTGCAACTAAACGCTTCTCACTG	1140				
Qy	1080	CCAGTACCCCTCAGCTCAGCAACCCCTTCTCCCATTTAACTCAAG	1139				
Db	1141	CCAGTACCCCTCAGCTCAGCAACCCCTTCTCCCATTTAACTCAAG	1200				

RESULT 3
SOYSBG7S 3393 bp DNA linear PLN 01-AUG-2002
LOCUS Glycine max gene for basic 7s globulin, complete cds.
DEFINITION Glycine max (soybean)
ACCESSION D16107
KEYWORDS Glycine max (soybean)
ORGANISM Glycine max (soybean)
Bukaryota; Viridiplantae; Streptophytina; Embryophyta; Tracheophyta; core eudicots; Spermatophytina; Magnoliophyta; Fabales; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1
AUTHORS Kagara H. and Hirano H.
TITLE Sequence of a cDNA encoding soybean basic 7s globulin
JOURNAL Nucleic Acids Res. 17 (21), 8868 (1989)
MEDLINE 90067863
PUBMED 2587227
REFERENCE 2
AUTHORS Watanabe Y. and Hirano H.
TITLE Nucleotide sequence of the basic 7s globulin gene from soybean
JOURNAL Plant Physiol. 105 (3), 1019-1020 (1994)
MEDLINE 94336768
PUBMED 8058830
REFERENCE 3
AUTHORS Watanabe Y.
TITLE Cloning and sequence analysis of the basic 7s globulin gene in soybean
JOURNAL Unpublished
REFERENCE 4
AUTHORS Watanabe Y.
TITLE Direct Submission
JOURNAL Submitted (23-APR-1993) Yoshihiro Watanabe, National Institute of Agrobiological Resources, Department of Molecular Biology, 2-1-2 Kannondai, Tsukuba, Ibaraki 305, Japan (Te: 81-908-38-7031, Fax: 81-908-38-7032)
FEATURES Location/Qualifiers
source 1..3393
/organism="Glycine max"
/mol_type="Genomic DNA"
/cultivar="Miyagishima"
/db_xref=taxon:38477
/clone="pGCBW2"
/tissue_type="leaf, etiolated"
/clone.Lib="pGCBW-Lib"
/dev_stage="Seedling"
misc_signal 1216..1230
/standard_name="heat shock like element"
/note="HSE"
/evidence-not_experimental
misc_feature 1254..1257
CAT_signal 1344..1350
TATA_signal 1369..1372
misc_feature 1369..1372
/note="primary transcription start site"

TITLE		Sequence and Analysis of Chromosome 2 of Dictyostelium	
JOURNAL		Unpublished	
REMARK		The Dictyostelium Genome Sequencing Consortium	
REFERENCE		2 (bases 1 to 56152)	
AUTHORS		Baumgart, C.	
DIRECT SUBMISSION		Submitted (04-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
COMMENT		CDS predictions from GenID may contain errors. Further Information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/)dictyostelium/ and the University Cologne, Institute for Biochemistry I (http://www.uni-koeln.de/dictyostelium/project.shtml)	
FUNDING		Deutsche Forschungsgemeinschaft (DFG).	
AGENCY		* NOTE: This is a 'working draft' sequence.	
* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.			
LOCATION/QUALIFIERS		1..56152 "organism="Dictyostelium discoideum"	
/mol-type="Genomic DNA"		/strain="Ax4"	
/db-xref="caxon:446689"		/chromosome="2"	
/map="#4790445..4846595"			
FEATURES		source	
FEATURES		lignin	
QUERY MATCH		6.0%; Score 82.4; DB 2; Length 56152;	
BEST LOCAL SIMILARITY		44.3%; Pred. No. 0_0011;	
MATCHES		430; Conservative 0; Mismatches 531; Indels 9; Gaps 2;	
DEFINITION		AC117076 linear - INV 12-MAR-2003	
DEFINITION		Dictyostelium discoideum chromosome 2 map	
DEFINITION		5862124..5045772 strain AX4, complete sequence.	
ACCESSION		AC117076	
VERSION		AC117076..2 GI:28828496	
SOURCE		HTG.	
ORGANISM		Dictyostelium discoideum	
ORGANISM		Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.	
REFERENCE		AC117076..1 (bases 1 to 183648)	
AUTHORS		Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P., Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K., Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and Noegel,A.A.	
TITLE		Sequence and analysis of chromosome 2 of Dictyostelium discoideum	
JOURNAL		Nature 418 (6893), 79-85 (2002)	
MEDLINE		22092622	
PUBMED		12097910	
REMARK		The Dictyostelium Genome Sequencing Consortium	
REMARK		2 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		3 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (06-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		4 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		5 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		6 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		7 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		8 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		9 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		10 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		11 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		12 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		13 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		14 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		15 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		16 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		17 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		18 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		19 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		20 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		21 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		22 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		23 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		24 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		25 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		26 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		27 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		28 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		29 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		30 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		31 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		32 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		33 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		34 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		35 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		36 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		37 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany	
REFERENCE		38 (bases 1 to 183648)	
AUTHORS		Baumgart,C.	
TITLE		Direct Submission	
TITLE		Submitted (04-MAR-2003) Genome Analysis	

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (02-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 4 (bases 1 to 1292401)

AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, A., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyns, S., Gordis, B., Grand-Pierre, N., Hagos, B., Horton, I., Huime, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Melorim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nenyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Patterson, K., Prunkhong, P., Pierre, N., Raymond, C., Retta, R., Ruse, C., Rogov, P., Roy, A., Schaefer, S., Schuback, R., Seaman, S., Severy, P., Smith, J.C., Spencer, B., Stojanovic, N., Stojanovic, N., Talamas, J., Testayre, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (30-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Jul 30, 2002 this sequence version replaced Gi:21306731.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1-997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web Site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: Li1383

Center clone name: 3118_D_11

FEATURES source

1. 1..12940

Location/Qualifiers

/clone lib="CTD-3118D11"

/clone lib="CTD2 Human BAC"

repeat_region 1..224

/rpt Family="AluSg/x"

repeat_region 225..403

/rpt Family="AluJb"

repeat_region 669..855

/rpt Family="LIMe3"

repeat_region 1519..1600

/rpt Family="AT rich"

repeat_region complement(1619..1895)

/rpt Family="AluBc"

repeat_region complement(1897..2197)

/rpt Family="AluSg"

repeat_region 2310..2593

/rpt Family="Charlie2"

repeat_region complement(2598..3451)

/rpt Family="LiNA4A"

unseq 2966..2969

/note="single clone coverage!"

repeat_region 3483..3713

/rpt Family="LIMda"

repeat_region complement(3714..4017)

/rpt Family="AluSg"

unseq 3739..3748

/note="<30 qual SNGI region"

repeat_region 380..391

/note="<30 qual SNGI region"

repeat_region 4018..4255

/rpt Family="LIMda"

repeat_region 4267..4982

/rpt Family="LIMD2"

repeat_region 4974..5464

/rpt Family="LIMD2"

repeat_region 5827..6015

/rpt Family="MIR"

repeat_region 8136..8345

/rpt Family="MIR"

repeat_region 8552..8737

/rpt Family="L2"

repeat_region 9347..9644

/rpt Family="AluSx"

repeat_region complement(10188..10324)

/rpt Family="MIR"

repeat_region 10170..11002

/rpt Family="MIR"

repeat_region 11116..11638

/rpt Family="MER41A"

repeat_region 12153..12222

/rpt Family="L12"

repeat_region complement(12468..12878)

/rpt Family="MIR"

repeat_region complement(13116..14221)

/rpt Family="MER41A"

repeat_region complement(15056..15245)

/rpt Family="AluJb"

repeat_region 15272..15714

/rpt Family="AluSx"

repeat_region complement(15162..16386)

/rpt Family="MIR"

repeat_region 16172..17008

/rpt Family="AluJb"

repeat_region 16930..19375

/rpt Family="AluJb"

repeat_region complement(18862..18963)

/rpt Family="MER5A"

repeat_region complement(19028..19089)

/rpt Family="L2"

repeat_region 19090..19375

/rpt Family="AluJb"

repeat_region complement(19376..19390)

/rpt Family="L12"

repeat_region complement(19480..19643)

/rpt Family="L2"

repeat_region complement(19744..19829)

/rpt Family="MIR3"

repeat_region complement(19976..20109)

/rpt Family="MIR"

repeat_region 22048..22074

/rpt Family="(T)n"

repeat_region complement(22240..22864)

/rpt Family="L1PA7"

repeat_region complement(23650..23559)

/rpt Family="AluJb"

repeat_region complement(24055..24374)

/rpt Family="AluSg"

repeat_region 24894..24922

/rpt Family="AT rich"

repeat_region complement(24934..25223)

/rpt Family="AluJb"

repeat_region 26058..26354

/rpt Family="AluSx"

repeat_region 26336..27229

/rpt Family="AluSx"

repeat_region 28099..28243

/rpt Family="AluSx"

repeat_region 29591..29611

Query Match 6.0% Score 81.6; DB 9; Length 129240;

Best Local Similarity 44.1% Pres. No. 0..0013;

Matches 383; Conservative 0; Mismatches 484; Indels 1; Gaps 1;

Bullerwell, C.E., Burger, G. and Lang, B.F. A novel motif for identifying rps3 homologs in fungal mitochondrial genomes							
AUTHORS							
TITLE							
JOURNAL							
Trends Biochem. Sci.							
25 (8), 363-365 (2000)							
MEDLINE							
10916154							
REFERENCE							
2 (bases 1 to 76568)							
Lang, B.F., O'Leary, C., Nerad, T., Gray, M.W. and Burger, G.							
AUTHORS							
TITLE							
The closest unicellular relatives of animals							
Curr. Biol.							
12 (20), 1773-1778 (2002)							
MEDLINE							
2288918							
PUBMED							
12401178							
REFERENCE							
3 (bases 64800 to 65249)							
Lang, F.B. and Bullerwell, L.C.							
AUTHORS							
TITLE							
Direct Submission							
JOURNAL							
Submitted (06-JUN-2000) Biochimie, Universite de Montreal, 2900							
Boul, Edouard Montréal, Québec H3C 3J4, Canada							
MEDLINE							
4 (bases 1 to 76568)							
AUTHORS							
Burger, G. and Lang, B.F.							
TITLE							
Direct Submission							
JOURNAL							
Submitted (10-APR-2002) Biochimie, Universite de Montreal, 2900							
Boul, Edouard Montréal, Québec H3T 1J4, Canada							
MEDLINE							
On Sep 29, 2002 this sequence version replaced gi:9966505.							
Location/Qualifiers							
1. . 76568							
organism="Monosiga brevicollis"							
/organism="mitochondrion"							
/mol_type="genomic DNA"							
/strain="ATCC 50154"							
/db_xref="ATCC:50154"							
COMMENT							
On Sep 29, 2002 this sequence version replaced gi:9966505.							
FEATURES							
source							
/product="Monosiga brevicollis"							
/organism="mitochondrion"							
/mol_type="genomic DNA"							
/strain="ATCC 50154"							
/db_xref="taxon:81824"							
/gene=rn1							
/gene=rn1							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							
/product="large subunit ribosomal RNA"							

/gene="rps8"
934. .8513
/gene="rps8"
/codon_start=1
/transl_table=4
/product="ribosomal protein 88"
/protein_id="PAN28337.1"
/db_xref="GI:23344063"
/translation="ANELISKINNNTYRMLTCQNMNDYSILKVLKSEGVIKLEI
NRNNKNKQTKTINPLDLGRKTYVNLKPKISPGKVKYIINKETISMSNTKKGTTINT
SKTGKTYLNSUHSEHKYLRTTSNNLQLINNKQITQEFAKNNVGGEILLVYTKFQ"
1059. .11287
/gene="atp6"
10529. .11287
/gene="atp6"
/CC number="3.6.1.34"
/codon_start=1
/transl_table=4
/product="ATP synthase F0 subunit 6"
/protein_id="PAN28338.1"
/db_xref="GI:23344062"
/translation="MAYSAFDOEFLIRIDPHLGMDISITNSTIEMFTALIVLIVL
YSGSGIKGSLSVPSRNOSVMEYEVFGVVEGRDKSQAAYFWLTVFMPFLVFF
NLFGIILTYTTEPSPAHATVTFGLSLSFIGTLLGMNGENYFSNEMPASSPLLAFF
MTVIELLISHTAKAGVSLGVRLAANITAGHLPAIISGFTWMILGGEMITASLPPMAI
11668. .11967
/gene="atp6"
11668. .11967
/EC number="3.6.1.34"
/codon_start=1
/transl_table=4
/product="ATP synthase F0 subunit 8"
/protein_id="PAN28339.1"
/db_xref="GI:23344063"
/translation="IMPOLSTIEFVSOYFWFLVILVYFLTISILVPIRAETIKTRKK
CMEDEESIWTISIDICIQTLANNENPAQOINVINNSVSNHINTIVSVSENANKWIK"
13729. .14760
/gene="nad1"
13729. .14760
/transl_table=4
/product="NADH dehydrogenase subunit 10"
/protein_id="PAN28340.1"
/db_xref="GI:23344064"
/translation="WTNLLITIETILKLLAVVPLINAVYLTAERKVLMQQLR
KGTAVALVFGLIQPLAGLKEAVLPTIANMIIIVVSEPAAPLILWATIYPN
ECKVIALDYNIGUYIFAVSSSYVAYTLMGSMANSKYYAFFAAIRAAOMISYESMGL
ILLSVLVCQGDNTITSIVLAQKINWIFIILPEPAINFSLAENINAPDILRGCEWE
ALKTGTEFLIVFVWTRATEPRFNYNTLMEWKSYLPSLIGFLTASILIAFNQYQI
SKENINLLY"
15143. .15214
/gene="rnt1"
15143. .15214
/product="tRNA-Thr"
15536. .15617
/gene="trns"
15536. .15617
/gene="trnS"
/product="tRNA-Ser"
15742. .15816
/gene="trnH"
15742. .15816
/gene="trnH"
/product="tRNA-His"
16105. .16500
/gene="rps12"
16105. .16500

/gene="rps12"
/codon_start=1
/transl_table=4
/product="ribosomal protein S12"
/protein_id="PAN28341.1"
/db_xref="GI:23344065"
/translation="MTTINQSLRKIRRKRTRITAGSKFLMNQGLINPQLKGVCIRVL
TRPKKPNSAQPKIAKYKLNLIVDAFIPEGHSLOQEHGIVLIRGRVKDLPGTRK
CVRKVDLAGVIGRKTRERSKYGTKKPTK"
17571. .18104
/gene="rpl16"
17571. .18104
/gene="rpl16"
/codon_start=1
/transl_table=4
/product="ribosomal protein L16"
/protein_id="PAN28342.1"
/db_xref="GI:23344066"
/translation="MLNKRKERTKSYSKYOKGRINSNLPYNSKLFQKNLENKDFRVL
IAMEPLRITASHAFAAGELAIREKLNKFSSKNSLSDTLEGKKNINGDYNLFVFPHL
DVNSKCGIKTKVVKVS"
19351. .19422
/gene="trnR"
19351. .19422
/gene="trnR"
/product="tRNA-Arg"
19369. .23450
/gene="trads"
join(19369. .20318, 22134. .23450)
/EC number="1.6.5.3"
/codon_start=1
/transl_table=4
/product="NADH dehydrogenase subunit 5"
/protein_id="PAN28343.1"
/db_xref="GI:23344067"
/translation="MYVLLFLPPLSVAIAGLGGFRFGTKGAGIITCIMTITSLSFW
TARVEGNMAMSPYTPLUKWTINGTYDMSLGIMYSLTVMEILYTWSCLVHLYST
SYNDGPPLPFRNSYLSLSPFTSRVMLVTFDLYSTFSSYKLDYTFNQYTRIQ
ANNSAIAKMMVTRVGDMLFLGMPAIPVFTNTLYDSTFSSYKLDYTFNQYTRIQ
ALCLTISLFLGAVGKSAQOLGHWTLDAMGSEPTPSALINAHMVTAGVULVLRSP
VALFHFLTHGPFVFLFQALFVQDLYKTAWSGSQLGYMVNLACGLSVM
TIEATACGAGAAATCGGTAAATATGTTAAAATAATTACCAACTAATTATACTTT
GPFPLAGYYSKDAILELAYGKTYTDGFHMLGLTAAAGMTAYYSTLILYMFPLSDTG

Query Match
Best Local Similarity 5.7% Score 77.4 ; DB 3 ; Length 76568;
Matches 382 ; Conservative 0 ; Mismatches 46 ; Indels 5 ; Gaps 2 ;

OY 4. AATATCTATACATTAATTAATTCATTGAAATTCTATGATTATTTTT 63
Db 71333 ATTATTATTATTAATTATTATTATTAATTATTATTATTAATTATT 71392
OY 64. TTITACCGAGAAATCGGTAAATATGTTAAAATAATTACCAACTAATTATACTTT 123
Db 71393 TTAAATTAAATTATTATTAATTATTATTATTAATTATTATTATTAATT 71451
OY 124. TAAGGCAATTAAGCATGTTGTTGATAAAATATATATTGTATAATCTTTCAAAGTA 183
Db 71452 AATTATTATTTAAATTAACCTTATTATTATTATTGTATTATTAATTATAATT 71511
OY 184. TAAAGTGTATAATGCGGTGTTGATATTGTTAGGTTCGATGCAAGTGTGCT 243
Db 71512 TAAGGATATTATAATATATCATATACTCTCTCTTTAAAGGGGAAATAATTAA 71571
OY 244. TTAGCATTTAGCCATTCTTCTTCTTACCAAAATAATGTAATGAAACCTTCTAG 303
Db 71572 GTATATTTGACATATTATTTTATTAATTATAATTATAATTATAATT 71631
OY 304. GAAAAAAGAAATGAAACATCATCCGGTGGAGTCGAGAACGCCAACCCAC 363
Db 71632 TTATATTATTAGTTAGTTAGTAAATAGAAATCTATTATAATTATAATTAA 71691

The remainder overlaps accession number AC091321
Li3165]. Location/qualifiers

```

FEATURES          source
repeat_region    /clone_id="RPCI-11 Human Male BAC"
                  /map="18"
                  /clone_start="794815"
                  /clone_end="794815"
repeat_region    27..70
repeat_region    /rpt_family="AT_rich"
repeat_region    528..682
repeat_region    /rpt_family="LIMBS5"
repeat_region    683..711
repeat_region    /rpt_family="LIMB"
repeat_region    712..1048
repeat_region    /rpt_family="LIMB5"
repeat_region    complement(1365..1726)
repeat_region    /rpt_family="LIn"
repeat_region    1770..1905
repeat_region    /rpt_family="C2A)n"
repeat_region    3047..3308
repeat_region    /rpt_family="MBR69A"
repeat_region    complement(2985)
repeat_region    /rpt_family="LIMB"
repeat_region    complement(3047..3308)
repeat_region    3734..3847
repeat_region    /rpt_family="LIMB"
repeat_region    complement(3848..4056)
repeat_region    /rpt_family="MTTID"
repeat_region    complement(4110..4183)
repeat_region    /rpt_family="MSD"
repeat_region    4184..4532
repeat_region    /rpt_family="MGSID"
repeat_region    complement(4799..4918)
repeat_region    /rpt_family="L1"
repeat_region    complement(5584..6121)
repeat_region    /rpt_family="ERYL"
repeat_region    complement(6193..6578)
repeat_region    /rpt_family="MTTII"
repeat_region    complement(6621..6745)
repeat_region    /rpt_family="MTTII-int"
repeat_region    6969..7174
repeat_region    /rpt_family="AluJO"
repeat_region    7575..7599
repeat_region    /rpt_family="AT_rich"
repeat_region    8207..8326
repeat_region    /rpt_family="TA)n"
repeat_region    8334..8357
repeat_region    /rpt_family="AT_rich"
repeat_region    8366..8393
repeat_region    /rpt_family="(TC)n"
repeat_region    complement(8432..8791)
repeat_region    /rpt_family="MTTII"
repeat_region    complement(8806..8839)
repeat_region    /rpt_family="ERYL"
repeat_region    8840..8860
repeat_region    /rpt_family="(CAA)n"
repeat_region    complement(8861..959)
repeat_region    /rpt_family="ERYL"
repeat_region    complement(9377..9799)
repeat_region    /rpt_family="MTTIA1"
repeat_region    complement(9993..10263)
repeat_region    /rpt_family="ERYL"
repeat_region    10316..10517
repeat_region    /rpt_family="AluX"
repeat_region    11315..11428
repeat_region    /rpt_family="MIR"
repeat_region    complement(13900..14227)
repeat_region    /rpt_family="AluSq"
repeat_region    complement(14647..14913)
repeat_region    /rpt_family="AluY"

```


Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and
 Barreil, B.
 Genome sequence of the human malaria parasite Plasmodium falciparum
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 SUBMITTED (02-OCT-2002) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 Location/Qualifiers
 1..11029
 /organism="Plasmodium falciparum 3D7"
 /mol_type="genomic DNA"
 /isolate="D7"
 /db_xref="taxon:363329"
 /chromosome="2"
 complement(1..53)
 /note="AT rich"
 /rpt_type=tandem
 repeat_region
 58..97
 /rpt_type=tandem
 complement(58..97)
 /rpt_type=tandem
 /rpt_unit="(TA)n"
 complement(117..156)
 /note="AT rich"
 /rpt_type=tandem
 repeat_region
 153..204
 /rpt_type=tandem
 complement(117..204)
 /rpt_type=tandem
 /rpt_unit="(TA)n"
 complement(241..264)
 /note="AT rich"
 /rpt_type=tandem
 complement(335..356)
 /rpt_type=tandem
 /rpt_unit="(A)n"
 complement(372..513)
 /note="AT rich"
 /rpt_type=tandem
 557..589
 /rpt_type=tandem
 complement(557..589)
 /rpt_type=tandem
 /rpt_unit="(TA)n"
 complement(659..684)
 /note="AT rich"
 /rpt_type=tandem
 709..844
 /rpt_type=tandem
 /rpt_unit="(TAAA)n"
 complement(916..961)
 /rpt_type=tandem
 /rpt_unit="(TAAA)n"
 complement(1035..1061)
 /rpt_type=tandem
 complement(1076..1122)
 /rpt_type=tandem
 /rpt_unit="(TA)n"
 complement(1294..1317)
 /rpt_type=tandem
 /rpt_unit="(TA)n"
 repeat_region
 1284..1313
 /rpt_type=tandem
 complement(1374..1672)
 /note="AT rich"
 /rpt_type=tandem
 repeat_region
 1931..1967
 /rpt_type=tandem
 complement(1122..1810)
 /note="AT rich"
 /rpt_type=tandem
 repeat_region
 4533..4618
 /rpt_type=tandem

Continuation (4 of 5) of AC116305 from base 300001 (AC116305 Dictyostelium discoideum)							
Query Match	Score	DB	Length	Start	End	Strand	Match/Sequence
Best Local Similarity	5.4%	3	Length 110000;				
Pred. No.	45.3%	2					
Matches	388	1	Mismatches 0;				
Conservative	459;	0	Indels 9;				
Gaps	3;						
Oy	7	AATTCTATCACTAAATTTGATTTAATTGTAGAAATTTCATGATTTPATTTTTTT	66	51285			
Db	51226	AATTCTATAAATACCTGACCGATTTTAAATTGCAGGTACAAAATAAACTTCGAAT	51285				
Qy	67	TACCAQAATCCGTTAATATTGTTAAATTACCAACTTAATTATAAAATTATTAA	126	51286			
Db	51286	AATTGATTTGAAATTATGATGCAAATCCTTTCAATTCTATAATCTGTATAATT	51345				
Qy	127	GGCAATTAAAGCATGTTGGAAAAATTATATATTGTTATAAAAT-----ACTTTTCAAAAGT	182				
Db	51346	ATTCTAAATGAAAGTCCTAAATATTCTCAAAATTATTAGTTTGCTATAATAATA	51405				
Qy	183	ATAAAGTTGATGATGGCGTGGTAGATTATTAGTTCTAGGTTCTGAACTGG 242		51406			
Db	51406	ATCTTATTATAATGATTTCTATTGTAATTATTAAATTATAATGATACATTC	51465				
Qy	243	TTTACACATTAGCTTATCTTTCTACCAAAATAATGTAATGGAAAACCTTA	302				
Db	51466	ATTCATCATAAACCTTTTAAATTGAAATTCAAACTCATCGCGTGGACTCTGAAAGGCCACCCA	51525				
Qy	303	GGAAAAAAAGAAATTCAAAATTGAAATTAAATAAGTGTCT -- TACAGTAAATGCCAATACCTTTT	419				
Db	51526	ATAAATAAAATTATAATTATAATTATAATTATAATTATAATAATAATAAT	51585				
Qy	363	CGTGACCCACAAATTAAATAAGTGTCT -- TACAGTAAATGCCAATACCTTTT	419				
Db	51586	ATTACCTTACCAAAAAATCAAGGATAAACTAAATTATGAACTTAAATTG	51645				
Qy	420	ATTCAAACTTTCCACTCTAAATTCTGGAGATTGGACCGTTAACATTAATAGTGT	479				
Db	51646	ATTAAATAATCACCACATTAAATTGAAATTCTGAAATTCAAACTAAACAGATA	51705				
Qy	480	TATATCCAAAGGTCTAAALAAARATTGTTACCTGTGCTTCACATTCAACTTTGCGAAC	539				
Db	51705	TTGAAAAAACTGTATGAATAAAATGACAACTAAAGTATCATGAACTMAAAACCATTTT	51765				
Qy	540	CTGAAAGCCGGTTAATGGTTAGTTGTTGCAAGTGTGCAAGTGTGCACTCAGGGC	599				
Db	51766	CAATTATTCTTCTTGGAAATTAAAGMTCTPAAACTAAAAACCTTCAAAATAAAA	51825				
Qy	600	TACCTGGTCTGGCTTGGCTCTTGTCAATT--TCACGTTGATTTGTGTGAACT	657				
Db	51826	TCTCTCTAAACCTGTTGTTGTTGTTGAACTATAAATCTGTAGTTGTTGATA	51885				
Qy	658	ACGGTACTGAAACTTATATAATTACATTAATTGCAATTTTAAAGCTTAAATAT	717				
Db	51886	TTCGAAACTCTGTGSCCAAACTTAAATTCTATATAATTAAATTGTTGGGTATA	51945				
Qy	718	ACTGATATAATATAATTGCGTTTAAATGCAATTTTAAAGCTTAAATATAAT	777				
Db	51946	TATCTTCTCTATATAATAAATATAAATATAAATATAAATATAAATATAAATATAA	52005				
Qy	778	AGATAAAATCGAAAAAATGTAAGGTTGAGGGGACAATAAAATGAAAG	837				
Db	52006	ATATAAATTAAATATAATAATAATAATAATAATAATAATAATAATAATAATA	52065				
Qy	838	TCTATTAAATTAAATTATAATAATAATAATAATAATAATAATAATAATAATA	853				

	FEATURES	SOURCE	The sequence of AC013666 has been incorporated into AC096550.
KEYWORDS	Homo sapiens (human)		Location/Qualifiers
SOURCE	Homo sapiens	1..101034	/organism="Homo sapiens"
ORGANISM	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		/mol_type="genomic DNA"
	Eukaryota; Metazoa; Primates; Catarrhini; Homidae; Homo;		/db_xref="Xaxon:9606"
REFERENCE	1 (bases 1 to 101034)		/chromosome="2"
AUTHORS	Sulston,J.E. and Waterston,R.		/map="2"
TITLE	Toward a complete human genome sequence		/clone="RP11-20F13"
JOURNAL	Genome Res. 8 (11). 1097-1108 (1998)		/clone.lib="RPCI-11"
MEDLINE	98470792		repeat_region
PUBLISHED	98470792	1..1075	/rpt_family="L1"
REFERENCE	2 (bases 1 to 101034)		/rpt_family="L1"
AUTHORS	Dauphin,S. and Meyer,R.	1376..1837	/rpt_family="L1"
TITLE	The sequence of Homo sapiens BAC clone RP11-20F13	1867..2087	/rpt_family="L1"
JOURNAL	Unpublished (2001)		/rpt_family="L1"
REFERENCE	3 (bases 1 to 101034)	1940..1982	/rpt_family="AT_rich"
AUTHORS	Waterston,R.H.		/rpt_family="AT_rich"
TITLE	Direct Submission	215..2209	/rpt_family="AT_rich"
JOURNAL	Submitted (18-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA	2197..2409	/rpt_family="AT_rich"
REFERENCE	4 (bases 1 to 101034)	2425..2447	/rpt_family="AT_rich"
AUTHORS	Waterston,R.	2429..3528	/rpt_family="AT_rich"
TITLE	Direct Submission	3535..3882	/rpt_family="AT_rich"
JOURNAL	Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA	4234..4285	/rpt_family="ALu"
COMMENT	On Mar 1, 2002 this sequence version replaced gi 1568715.	4240..4283	/rpt_family="ERV1"
	-- Genome Center	4508..4628	/rpt_family="TG1n"
	Center: Washington University Genome Sequencing Center	4525..5660	/rpt_family="TG1n"
	Center Code: WUGSC	5665..5928	/rpt_family="MIR"
	Web site: http://genome.wustl.edu/gsc	6002..6091	/rpt_family="MIR"
	Contact: sapiens@genome.wustl.edu	8205..8279	/rpt_family="MIR"
	-- Summary Statistics	8324..9069	/rpt_family="SERV1"
	Center project name: H_NH0020F13	11481..11501	/rpt_family="L1"
	Drafting Center: WIBR	1231..12573	/rpt_family="AT_rich"
	--	12674..12864	/rpt_family="SERV1"
	--	14233..14333	/rpt_family="L1"
	--	14987..15130	/rpt_family="AT_rich"
	--	15471..15876	/rpt_family="SERV1"
	--	17889..17889	/rpt_family="AT_rich"
	--	17904..18466	/rpt_family="SERV1"
	--	18122..18142	/rpt_family="AT_rich"
	--	18455..18517	/rpt_family="AT_rich"
	--	18538..18865	/rpt_family="MIR2_type"
	--	18866..18920	/rpt_family="AT_rich"
	--	18903..19656	/rpt_family="AT_rich"
	--	20507..20646	/rpt_family="L1"
	NEIGHBORING SEQUENCE INFORMATION:		repeat_region
	The RP11-11 human BAC library was made from the blood of one male donor, as described by Ossegawa,K., Woon,P.Y., Zhao,B., Frangen,E., McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc		repeat_region
	SOURCE INFORMATION:		repeat_region
	Mapping information for this clone was provided by Dr. John D. Tatteno, M., Catanese,J.J. and de Jong, P.J. (1998) An improved approach to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone, and the assembly was confirmed by restriction digest.		repeat_region
	The RP11-11 human BAC library was made from the blood of one male donor, as described by Ossegawa,K., Woon,P.Y., Zhao,B., Frangen,E., McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc		repeat_region
	SOURCE INFORMATION:		repeat_region
	The clone sequenced to the right is RP11-33SE8, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-20F13; actual end is at base position 70865 of RP11-33SE8.		repeat_region
	Data from AC083850 was used to finish this clone, AC013666.		repeat_region

BEST AVAILABLE COPY